Neighborhood filters and PDE's

Antoni Buades, Bartomeu Coll, Jean-Michel Morel

▶ **To cite this version:**

Antoni Buades, Bartomeu Coll, Jean-Michel Morel. Neighborhood filters and PDE's. Numerische Mathematik, 2006, 105 (1), pp.1-34. hal-00271142

HAL Id: hal-00271142
https://hal.science/hal-00271142

Submitted on 21 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Neighborhood filters and PDE's

Antoni Buades*† Bartomeu Coll * Jean-Michel Morel †‡

Abstract

Denoising images can be achieved by a spatial averaging of nearby pixels. However, although this method removes noise it creates blur. Hence, neighborhood filters are usually preferred. These filters perform an average of neighboring pixels, but only under the condition that their grey level is close enough to the one of the pixel in restoration. This very popular method unfortunately creates shocks and staircasing effects. In this paper, we perform an asymptotic analysis of neighborhood filters as the size of the neighborhood shrinks to zero. We prove that these filters are asymptotically equivalent to the Perona-Malik equation, one of the first nonlinear PDE's proposed for image restoration. As a solution, we propose an extremely simple variant of the neighborhood filter using a linear regression instead of an average. By analyzing its subjacent PDE, we prove that this variant does not create shocks: it is actually related to the mean curvature motion. We extend the study to more general local polynomial estimates of the image in a grey level neighborhood and introduce two new fourth order evolution equations.

1 Introduction

According to Shannon's theory, a signal can be correctly represented by a discrete set of values, the "samples", only if it has been previously smoothed. Let us start with u_0 the physical image, a real function defined on a bounded domain $\Omega \subset \mathbb{R}^2$. Then a blur optical kernel k is applied, i.e. u_0 is convolved with k to obtain an observable signal $k * u_0$. Gabor remarked in 1960 that the difference between the original and the blurred images is roughly proportional to its Laplacian, $\Delta u = u_{xx} + u_{yy}$. In order to formalize this remark, we have to notice that k is spatially concentrated, and that we may introduce a scale parameter for k, namely $k_h(\mathbf{x}) = h^{-1}k(h^{-\frac{1}{2}}\mathbf{x})$. If, for instance, u is C^2 and bounded and if k is a radial function in the Schwartz class, then

$$\frac{u_0 * k_h(\mathbf{x}) - u_0(\mathbf{x})}{h} \rightarrow c\Delta u_0(\mathbf{x}).$$

Hence, when h gets smaller, the blur process looks more and more like the heat equation

$$u_t = c\Delta u, \quad u(0) = u_0.$$

Thus, Gabor established a first relationship between local smoothing operators and PDE's. The classical choice for k is the Gaussian $G_h(\mathbf{x}) = \frac{1}{(4\pi h^2)} e^{-\frac{|\mathbf{x}|^2}{4h^2}}$. Such a convolution blurs the discontinuities

*University of Balearic Islands, Ctra. Valldemossa Km. 7.5, 07122 Palma de Mallorca, Spain (e-mail: tomeu.coll@uib.es)

†J.M Morel is with the CMLA, ENS Cachan 61, Av du President Wilson 94235 Cachan, France (e-mail: buades@cmla.ens-cachan.fr, morel@cmla.ens-cachan.fr)

‡This work has been partially financed by the Centre National d'Etudes Spatiales (CNES), the Office of Naval Research under grant N00014-97-1-0839, the Ministerio de Ciencia y Tecnologia under grant MTM2005-08567. During this work, the first author had a fellowship of the Govern de les Illes Balears for the realization of his PhD.

and removes the high frequency features of the image. For this reason, more recently introduced filters try to adapt to the local configuration of the image.

The anisotropic filter (*AF*) attempts to avoid the blurring effect of the Gaussian by *convolving the image u at $\boldsymbol{x}$ only in the direction orthogonal to the gradient $Du(\boldsymbol{x})$*. If $Du(\mathbf{x}) \neq 0$, let us denote the tangent and orthogonal directions to the level line passing through $\mathbf{x}$ by $\xi = Du(\mathbf{x})^{\perp}/|Du(\mathbf{x})|$ and $\eta = Du(\mathbf{x})/|Du(\mathbf{x})|$, respectively. Then,

$$AF_h u(\mathbf{x}) = \int G_h(t)\, u(\mathbf{x} + t\xi) dt,$$

where $G_h(t) = \frac{1}{\sqrt{2\pi}h} e^{-\frac{t^2}{2h^2}}$ denotes the one-dimensional Gauss function with variance h^2. At points where $Du(\mathbf{x}) = 0$ an isotropic Gaussian mean is applied. The anisotropic filter better preserves discontinuities, but performs poorly on flat and textured regions.

Neighborhood filters are based on the idea that all pixels belonging to the same object have a similar grey level value. The neighborhood filters [15, 28] therefore take an average of the values of pixels which are both close in grey level value and spatial distance. For $\mathbf{x} \in \Omega$, define

$$YNF_{h,\rho} u(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{B_\rho(\mathbf{x})} u(\mathbf{y}) e^{-\frac{|u(\mathbf{y})-u(\mathbf{x})|^2}{h^2}} d\mathbf{y}, \tag{1}$$

where $B_\rho(\mathbf{x})$ is a ball of center $\mathbf{x}$ and radius ρ, h is the filtering parameter and $C(\mathbf{x}) = \int_{B_\rho(\mathbf{x})} e^{-\frac{|u(\mathbf{y})-u(\mathbf{x})|^2}{h^2}} d\mathbf{y}$ is the normalization factor. The Yaroslavsky filter (1) is less known than more recent versions, namely the *SUSAN filter* [25] and the *Bilateral filter* [26]. Both algorithms, instead of considering a fixed spatial neighborhood $B_\rho(\mathbf{x})$, weigh the distance to the reference pixel $\mathbf{x}$,

$$SNF_{h,\rho} u(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{\Omega} u(\mathbf{y}) e^{-\frac{|\mathbf{y}-\mathbf{x}|^2}{\rho^2}} e^{-\frac{|u(\mathbf{y})-u(\mathbf{x})|^2}{h^2}} d\mathbf{y}, \tag{2}$$

where $C(\mathbf{x}) = \int_{\Omega} e^{-\frac{|\mathbf{y}-\mathbf{x}|^2}{\rho^2}} e^{-\frac{|u(\mathbf{y})-u(\mathbf{x})|^2}{h^2}} d\mathbf{y}$ is the normalization factor and ρ is now a spatial filtering parameter. In practice there is no serious difference between $YNF_{h,\rho}$ and $SNF_{h,\rho}$. The performance of both algorithms is justified by the same arguments. Inside a homogeneous region, the grey level values slightly fluctuate because of the noise. In this case, the first strategy computes an arithmetic mean of the neighborhood and the second strategy a Gaussian mean. At a contrasted edge separating two regions, if the grey level difference between both regions is larger than h, both algorithms compute averages of pixels belonging to the same region as the reference pixel. Thus, the algorithm does not blur the edges, which is its main scope.

The main objective of this paper is to apply the Gabor method to the neighborhood filters. We shall compute the subjacent PDE of the neighborhood filter. This leads to a comparison of this PDE with another well-known PDE in image filtering, the Perona-Malik equation. Thanks to this comparison, some artifacts of neighborhood filters, shocks and staircase effects, will be explained.

This will lead us to propose a slight modification of the neighborhood filters, replacing the average by a linear regression. Studying the asymptotic expansion of this new filter by the Gabor method again, we will prove that linear regression has a well posed subjacent PDE, namely a mean curvature motion.

Our plan is as follows: in section 2, we review the PDE models in image processing. In section 3 we perform, for the sake of clarity, the asymptotic analysis of neighborhood filters in dimension 1. Section 4 is dedicated to N dimensions. Section 5 introduces a new neighborhood filter using a linear regression. This filter creates no artefact. In section 6 we extend the same ideas to the vector valued case. In a final extension (section 7), we examine the application of the method to more general polynomial local interpolations.

2 PDE based models

Remarking that the optical blur is equivalent to one step of the heat equation, Gabor deduced that we can, to some extent, deblur an image by reversing the time in the heat equation, $u_t = -\Delta u$. Numerically, this amounts to subtracting the filtered version from the original image,

$$u - G_h * u = -h^2 \Delta u + o(h^2).$$

This leads to considering the reverse heat equation as an image restoration, ill-posed though it is. The time-reversed heat equation was stabilized in the Osher-Rudin shock filter [17] who proposed

$$u_t = -\text{sign}(\mathcal{L}(u))|Du|, \tag{3}$$

where the propagation term $|Du|$ is tuned by the sign of an edge detector $\mathcal{L}(u)$. The function $\mathcal{L}(u)$ changes sign across the edges where the sharpening effect therefore occurs. In practice, $\mathcal{L}(u) = \Delta u$ and the equation is related to an inverse heat equation.

The Osher-Rudin equation is equivalent to the so-called Kramer filter proposed in the seventies [14]. This filter replaces the gray level value at a point by either the minimum or maximum of the gray level values in a neighborhood. This choice depends on which is closest to the current value, in the same spirit as the neighborhood filters. Let us use KF_h to denote the Kramer operator where h stands for the radius of the neighborhood. Schavemaker et al. [23] showed that in one dimension the Kramer filter is equivalent to the Osher-Rudin shock filter,

$$KF_h u - u = -\text{sign}(u'')\,|u'|\,h^2 + o(h^2).$$

In the two dimensional case, Guichard et al. [7] showed that the Laplacian must be replaced by a directional second derivative of the image $D^2u(Du, Du)$,

$$KF_h u - u = -\text{sign}(D^2u(Du, Du))\,|Du|\,h^2 + o(h^2).$$

The early Perona-Malik "anisotropic diffusion" [18] is directly inspired from the Gabor remark. It reads

$$u_t = \text{div}(g(|Du|^2)Du), \tag{4}$$

where $g : [0, +\infty) \to [0, +\infty)$ is a smooth decreasing function satisfying $g(0) = 1$, $\lim_{s\to+\infty} g(s) = 0$. This model is actually related to the preceding ones. Let us consider the second derivatives of u in the directions of Du and $Du^\perp$,

$$u_{\eta\eta} = D^2u(\frac{Du}{|Du|}, \frac{Du}{|Du|}), \quad u_{\xi\xi} = D^2u(\frac{Du^\perp}{|Du|}, \frac{Du^\perp}{|Du|}).$$

Then, the equation (4) can be rewritten as

$$u_t = g(|Du|^2)u_{\xi\xi} + h(|Du|^2)u_{\eta\eta}, \tag{5}$$

where $h(s) = g(s) + 2sg'(s)$. Perona and Malik proposed the function $g(s) = \frac{1}{1+s/k}$. In this case, the coefficient of the first term is always positive and this term therefore appears as a one dimensional diffusion term in the orthogonal direction to the gradient. The sign of the second coefficient, however, depends on the value of the gradient. When $|Du|^2 < k$ this second term appears as a one dimensional diffusion in the gradient direction. It leads to a reverse heat equation term when $|Du|^2 > k$.

In [6] and earlier in [9], serious mathematical attempts were made to define a solution to the Perona-Malik equation beyond blow-up. Kichenassamy defined entropic conditions across the singularities. Esedoglu analyzes the asymptotic behavior of the numerical scheme proposed by Perona and Malik in

the 1D case and proves that it obeys a system of partial differential equations with special boundary conditions on the singularities. These analyses substantiate a staircasing effect, as was indeed observed in the experiments of the present paper. The analysis by Esedoglu suggests that the solution eventually becomes constant. Indeed, he proves that flat regions tend to merge at a speed depending upon their contrast with the neighboring ones. Then, the whole process boils down to a fine to coarse image segmentation process. Unfortunately, to the best of our knowledge, no similar analysis is available in 2D.

The Perona-Malik model has got many variants and extensions. Tannenbaum and Zucker [10] proposed, endowed in a more general shape analysis framework, the simplest equation of the list,

$$u_t = |Du| \mathrm{div}\left(\frac{Du}{|Du|}\right) = u_{\xi\xi}.$$

This equation had been proposed some time before in another context by Sethian [24] as a tool for front propagation algorithms. This equation is a "pure" diffusion in the direction orthogonal to the gradient and is equivalent to the anisotropic filter *AF*, since

$$AF_h u - u = \frac{1}{2} u_{\xi\xi} h^2 + o(h^2).$$

This diffusion is also related to two models proposed in image restoration. The Rudin-Osher-Fatemi [19] total variation model leads to the minimization of the total variation of the image $TV(u) = \int |Du|$, subject to some constraints. The steepest descent of this energy reads, at least formally,

$$\frac{\partial u}{\partial t} = \mathrm{div}\left(\frac{Du}{|Du|}\right) \tag{6}$$

which is related to the mean curvature motion and to the Perona-Malik equation when $g(|Du|^2) = \frac{1}{|Du|}$. This particular case, which is not considered in [18], yields again (6). An existence and uniqueness theory is available for this equation [2].

3 Asymptotic behavior of neighborhood filters (dimension 1)

Let u denote a one-dimensional signal defined on an interval $I \subset \mathbb{R}$ and consider the neighborhood filter

$$YNF_{h,\rho} u(x) = \frac{1}{C(x)} \int_{x-\rho}^{x+\rho} u(y) e^{-\frac{|u(y)-u(x)|^2}{h^2}} dy, \tag{7}$$

where $C(x) = \int_{x-\rho}^{x+\rho} e^{-\frac{|u(y)-u(x)|^2}{h^2}} dy$.

Theorem 3.1 *Suppose $u \in C^2(I)$, and let $\rho, h, \alpha > 0$ such that $\rho, h \to 0$ and $h = O(\rho^\alpha)$. Consider the continuous function $g(t) = \frac{te^{-t^2}}{E(t)}$, for $t \neq 0$, $g(0) = \frac{1}{2}$, where $E(t) = 2\int_0^t e^{-s^2}\, ds$. Let f be the continuous function*

$$f(t) = \frac{g(t)}{t^2} + g(t) - \frac{1}{2t^2}, \quad f(0) = \frac{1}{6}.$$

Then, for $x \in \mathbb{R}$,

1. *If $\alpha < 1$, $\quad YNF_{h,\rho} u(x) - u(x) \simeq \frac{u''(x)}{6}\rho^2$.*
2. *If $\alpha = 1$, $\quad YNF_{h,\rho} u(x) - u(x) \simeq f(\frac{\rho}{h}|u'(x)|)\, u''(x)\, \rho^2$.*
3. *If $1 < \alpha < \frac{3}{2}$, $\quad YNF_{h,\rho} u(x) - u(x) \simeq g(\rho^{1-\alpha}|u'(x)|)\, u''(x)\, \rho^2$.*

Proof: First, we rewrite the difference $YNF_{h,\rho}u(x) - u(x)$ as

$$YNF_{h,\rho}u(x) - u(x) = \frac{1}{C(x)} \int_{-\rho}^{\rho} (u(x+t) - u(x)) e^{-\frac{|u(x+t)-u(x)|^2}{h^2}}\, dt.$$

Taking the Taylor expansion of $u(x+t)$ for $t \in (-\rho, \rho)$,

$$YNF_{h,\rho}u(x) - u(x) = \frac{1}{C(x)} \int_{-\rho}^{\rho} (u't + u''\frac{t^2}{2} + O(t^3)) e^{-(\frac{u'^2t^2}{h^2} + \frac{u'u''t^3}{h^2} + O(\frac{t^4}{h^2}))}\, dt, \tag{8}$$

where $C(x) = \int_{-\rho}^{\rho} e^{-(\frac{u'^2t^2}{h^2} + \frac{u'u''t^3}{h^2} + O(\frac{t^4}{h^2}))}\, dt$. If $\alpha < 1$, $\frac{\rho^2}{h^2}$ tends to zero and we can expand the exponential function in (8),

$$\begin{aligned} YNF_{h,\rho}u(x) - u(x) &= \frac{1}{C(x)} \int_{-\rho}^{\rho} (u't + u''\frac{t^2}{2} + O(t^3))(1 - \frac{u'^2t^2}{h^2} - \frac{u'u''t^3}{h^2} + O(\frac{t^4}{h^2}))\, dt \\ &\simeq \frac{1}{2\rho}\frac{u''\rho^3}{3}. \end{aligned}$$

This proves *(1)*. If $1 \leq \alpha < \frac{3}{2}$, we cannot apply the above expansion. However, $\frac{\rho^3}{h^2} \to 0$, and we can decompose the exponential as $e^{-\frac{|u(x+t)-u(t)|^2}{h^2}} = e^{\frac{-u'^2t^2}{h^2}}(1 - \frac{u'u''t^3}{h^2} + O(\frac{t^4}{h^2}))$. Then, we approximate the difference (8) by

$$\begin{aligned} YNF_{h,\rho}u(x) - u(x) &= \frac{1}{C(x)} \int_{-\rho}^{\rho} (u't + \frac{u''t^2}{2} + O(t^3))\, e^{\frac{-u'^2t^2}{h^2}}(1 - \frac{u'u''t^3}{h^2} + O(\frac{t^4}{h^2}))\, dt \\ &\simeq \left[\frac{e^{-\frac{\rho^2u'^2}{h^2}}}{E(\frac{\rho}{h}|u'|)}\frac{h}{\rho u'} + \frac{e^{-\frac{\rho^2u'^2}{h^2}}}{E(\frac{\rho}{h}|u'|)}\frac{\rho u'}{h} - \frac{h^2}{2\rho^2u'^2} \right] u''\rho^2 \end{aligned}$$

If $h \simeq \rho$, all the terms in the above expression have the same order ρ^2 and using the definitions of f and g we obtain *(2)*. Finally, when $h \simeq \rho^\alpha$, $1 < \alpha < \frac{3}{2}$, we keep the lower order term and obtain *(3)*. □

Interpretation

According to Theorem 3.1, the neighborhood filter makes the signal evolve proportionally to its second derivative. The equation $u_t = cu''$ acts as a smoothing or enhancing model depending on the sign of c. Following the previous theorem, we can distinguish three cases depending on the values of h and ρ. First, if h is much larger than ρ the second derivative is weighted by a positive constant and the signal is therefore filtered by a heat equation. Second, if h and ρ have the same order, the sign and magnitude of the weight is given by $f(\frac{\rho}{h}|u'(x)|)$. As the function f takes positive and negative values (see Figure 4), the filter behaves as a filtering/enhancing algorithm depending on the magnitude of $|u'(x)|$. If B denotes the zero of f, then a filtering model is applied wherever $|u'| < B\frac{h}{\rho}$ and an enhancing model wherever $|u'| > B\frac{h}{\rho}$. The intensity of the enhancement tends to zero when the derivative tends to infinity. Thus, points x where $|u'(x)|$ is large are not altered. The transition of the filtering to the enhancement model creates a singularity in the filtered signal. In the last case, ρ is much larger than h and the sign and magnitude of the weight is given by $g(\frac{\rho}{h}|u'(x)|)$. Function g is positive and decreases to zero. If the derivative of u is bounded then $\frac{\rho}{h}|u'(x)|$ tends to infinity and the intensity of the filtering to zero. In this case, the signal is hardly modified.

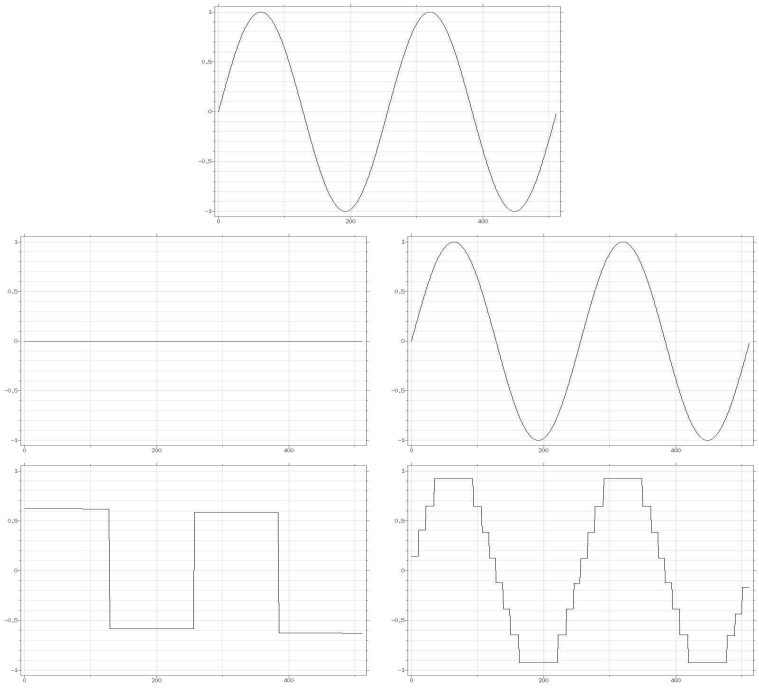

Figure 1: One dimensional neighborhood filter experiment. The neighborhood filter is iterated until the steady state is attained for different values of the ratio ρ/h. Top: Original sine signal. Middle left: filtered signal with $\rho/h = 10^{-8}$. Middle right: filtered signal with $\rho/h = 10^{8}$. Bottom left: filtered signal with $\rho/h = 2$. Bottom right: filtered signal with $\rho/h = 5$. The examples corroborate the results of Theorem 3.1. If ρ/h tends to zero the algorithm behaves like a heat equation and the filtered signal tends to a constant. If, instead, ρ/h tends to infinity the signal is hardly modified. If ρ and h have the same order, the algorithm presents a filtering/enhancing dynamic. Singularities are created due to the transition of smoothing to enhancement. The number of enhanced regions strongly depends upon the ratio $\frac{\rho}{h}$ as illustrated in the bottom figures.

In summary, a neighborhood filter in dimension 1 shows interesting behavior only if ρ and h have the same order of magnitude; in which case the neighborhood filter behaves like a Perona-Malik equation. It enhances edges with a gradient above a certain threshold and smoothes the rest.

Figure 1 illustrates the behavior of the one dimensional neighborhood filter. The algorithm is iterated until the steady state is attained on a sine signal for different values of the ratio ρ/h. The results of the experiment corroborate the asymptotical expansion of Theorem 3.1. In the first experiment, $\rho/h = 10^{-8}$ and the neighborhood filter is equivalent to a heat equation. The filtered signal tends to a constant. In the second experiment, $\rho/h = 10^{8}$ and the value $g(\frac{\rho}{h}|u'|)$ is nearly zero. As predicted by the theorem, the filtered signal is nearly identical to the original one. The last two experiments illustrate the filtering/enhancing behavior of the algorithm when h and ρ have similar values. As predicted, an enhancing model is applied where the derivative is large. Many singularities are being created because of the transition of the filtering to the enhancing model. Unfortunately, the number of singularities and their position depend upon the value of ρ/h. This behavior is explained by Theorem 3.1(2). Figure 8 illustrates the same effect in the 2D case.

The long term consistency of an iterated neighborhood filter with the Perona-Malik equation is difficult to explore numerically. We must indeed distinguish the theoretical Perona-Malik equation

from its numerical schemes. All finite diference schemes create some numerical diffusion. Thus, they are not exactly consistent with the Perona-Malik equation. This explains why iterating such schemes can yield an asymptotic constant steady state. The Perona Malik equation is ill-posed and no existence-uniqueness theory is available to the best of our knowledge [1]. Now, it seems sound to conjecture that piecewise constant functions with smooth jumps could be steady states to the equation in a suitable existence theory. It is easy to check that such functions are also steady states to an iterated neighborhood filter, provided the scale parameter is small enough. In summary, neighborhood filters yield an implementation of the Perona-Malik consistent with piecewise constant steady states.

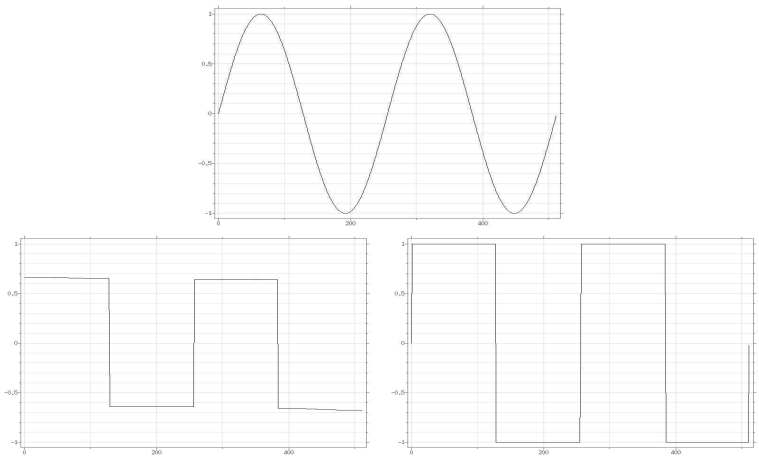

Figure 2: Comparison between the neighborhood filter and the shock filter. Top: Original signal. Bottom left: application of the neighborhood filter. Bottom right: application of the shock filter. The minimum and the maximum of the signal have been preserved by the shock filter and reduced by the neighborhood filter. This fact illustrates the filtering/enhancing character of the neighborhood filter compared with a pure enhancing filter.

The filtering/enhancing character of the neighborhood filter is very different from a pure enhancing algorithm like the Osher-Rudin shock filter. Figures 2 and 3 illustrate these differences. In Figure 2, the minimum and the maximum of the signal have been preserved by the shock filter, while these two values have been significantly reduced by the neighborhood filter. This filtering/enhancing effect is optimal when the signal is noisy. Figure 3 shows how the shock filter creates artificial steps due to the fluctuations of noise, while the neighborhood filter reduces the noise avoiding any spurious shock. Parameter h has been chosen larger than the amplitude of noise in order to remove it. Choosing an intermediate value of h, artificial steps could also be generated on points where the noise amplitude is above this parameter value.

4 The N dimensional case

The previous theorem can be extended to the N-dimensional case, with the 2D and 3D case being the most interesting for image processing purposes. Films can also be modeled by a 3D spatiotemporal space.

Let $u(x_1, \ldots, x_N)$ be defined on a bounded domain $\Omega \subset \mathbb{R}^N$ and $\mathbf{x} \in \Omega$. Assume that $Du(\mathbf{x}) \neq 0$ and let us denote $e_1 = Du(\mathbf{x})/|Du(\mathbf{x})|$. Let $e_2, \ldots, e_N$ be an orthonormal basis of the hyperplane orthogonal to Du, $Du^{\perp}$. Then the set $\{e_1, e_2, \ldots, e_N\}$ forms an orthonormal basis of $\mathbb{R}^N$.

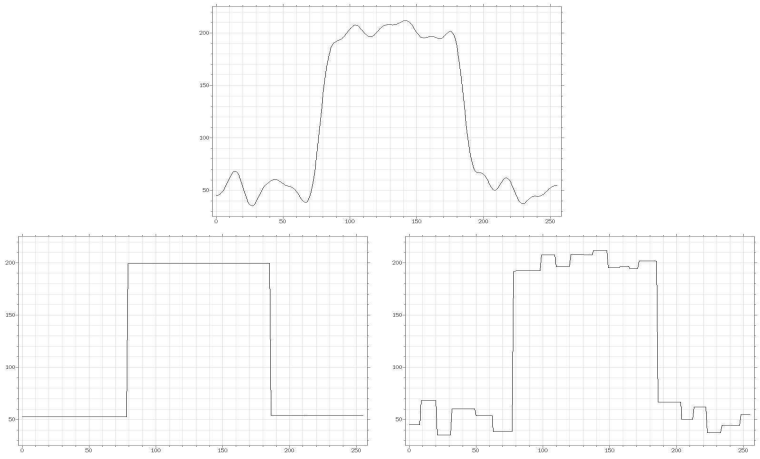

Figure 3: Comparison between the neighborhood filter and the shock filter. Top: Original signal. Bottom left: application of the neighborhood filter. Bottom right: application of the shock filter. The shock filter is sensitive to noise and creates spurious steps. The filtering/enhancing character of the neighborhood filter avoids this effect.

Theorem 4.1 *Let $u \in C^2(\Omega)$ and $\rho, h, \alpha > 0$ such that $\rho, h \to 0$ and $h = O(\rho^\alpha)$. Let us consider the continuous function $\tilde{g}$ defined by $\tilde{g}(t) = \frac{1}{3}\frac{te^{-t^2}}{E(t)}$, for $t \neq 0$, $\tilde{g}(0) = \frac{1}{6}$, where $E(t) = 2\int_0^t e^{-s^2}\, ds$. Let $\tilde{f}$ be the continuous function defined as*

$$\tilde{f}(t) = 3\tilde{g}(t) + \frac{3\tilde{g}(t)}{t^2} - \frac{1}{2t^2}, \quad \tilde{f}(0) = \frac{1}{6}.$$

Then, for $\boldsymbol{x} \in \Omega$,

1. *If $\alpha < 1$,*

$$YNF_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq \frac{\triangle u(\boldsymbol{x})}{6}\rho^2.$$

2. *If $\alpha = 1$,*

$$\begin{aligned} YNF_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \quad \simeq \quad & \Big[\tilde{f}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)\; D^2u(e_1,e_1)(\boldsymbol{x}) + \\ & \tilde{g}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)\; \sum_{i=2}^{N} D^2u(e_i,e_i)(\boldsymbol{x})\Big]\rho^2 \end{aligned}$$

3. *If $1 < \alpha < \frac{3}{2}$,*

$$YNF_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq \tilde{g}(\rho^{1-\alpha}|Du(\boldsymbol{x})|)\left[\, 3\, D^2u(e_1,e_1)(\boldsymbol{x}) + \sum_{i=2}^{N} D^2u(e_i,e_i)(\boldsymbol{x})\right]\rho^2$$

Proof: First, we rewrite the difference $YNF_{h,\rho}u(\mathbf{x}) - u(\mathbf{x})$ as

$$YNF_{h,\rho}u(\mathbf{x}) - u(\mathbf{x}) = \frac{1}{C(\mathbf{x})}\int_{B_\rho(0)} (u(\mathbf{x}+\mathbf{t}) - u(\mathbf{x}))e^{-\frac{|u(\mathbf{x}+\mathbf{t})-u(\mathbf{x})|^2}{h^2}}\, d\mathbf{t}.$$

Taking the Taylor expansion of u$(\mathbf{x}+\mathbf{t})$ for $\mathbf{t} \in B_\rho(0)$,

$$u(\mathbf{x}+\mathbf{t}) = u(\mathbf{x}) + pt_1 + \sum_{i,j\in\{1,\dots,N\}} q_{ij}t_it_j + O(|\mathbf{t}|^3),$$

where $\mathbf{t} = (t_1, \dots, t_N)$, $p = |Du(\mathbf{x})|$ and if $p > 0$,

$$q_{ii} = \frac{1}{2}D^2u(e_i, e_i)(\mathbf{x}), \quad q_{ij} = D^2u(e_i, e_j)(\mathbf{x}) \quad \text{if} \quad i \neq j.$$

When $\alpha < 1$, we expand the exponential function and obtain

$$\begin{aligned} YNF_{h,\rho}u(\mathbf{x}) - u(\mathbf{x}) &\simeq \frac{1}{C(\mathbf{x})}\int_{B_\rho(0)} (pt_1 + \sum_{i,j} q_{ij}t_it_j)(1 - \frac{p^2t_1^2}{h^2})\, d\mathbf{t} \\ &\simeq \frac{1}{4\rho^2}\frac{2\triangle(\mathbf{x})u\rho^4}{3}. \end{aligned}$$

This proves *(1)*. When $1 \leq \alpha < \frac{3}{2}$, we cannot apply the above expansion because $\frac{\rho^2}{h^2}$ does not tend to zero. However, $\frac{\rho^3}{h^2} \to 0$, and we can decompose the exponential as

$$e^{-\frac{|u(\mathbf{x}+\mathbf{t})-u(\mathbf{x})|^2}{h^2}} = e^{\frac{-p^2t_1^2}{h^2}}(1 - \frac{2pt_1}{h^2}\sum_{i,j} q_{ij}t_it_j + O(\frac{|t|^4}{h^2})).$$

Using the Taylor expansion of u and of the above exponential function we obtain

$$\begin{aligned} YNF_{h,\rho}u(\mathbf{x}) - u(\mathbf{x}) &\simeq \frac{1}{C(\mathbf{x})}\left(\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2\, d\mathbf{t} - \frac{2p^2}{h^2}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^4\, d\mathbf{t}\right) q_{11} \\ &\quad + \frac{1}{C(x)}\sum_{i=2}^{N}\left(\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_i^2\, d\mathbf{t} - \frac{2p^2}{h^2}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2t_i^2\, d\mathbf{t}\right) q_{ii} \end{aligned}$$

where $C(\mathbf{x}) \simeq \int e^{\frac{-p^2t_1^2}{h^2}} d\mathbf{t}$. Now, we compute the previous integrals and obtain

$$YNF_{h,\rho}u(\mathbf{x}) - u(\mathbf{x}) \simeq \left(\frac{2e^{\frac{-p^2t_1^2}{h^2}}}{E(\frac{\rho p}{h})}\frac{\rho p}{h} + \frac{2e^{\frac{-p^2t_1^2}{h^2}}}{E(\frac{\rho p}{h})}\frac{h}{\rho p} - \frac{h^2}{\rho^2p^2}\right) q_{11}\rho^2 + \frac{2e^{\frac{-p^2t_1^2}{h^2}}}{3E(\frac{\rho p}{h})}\frac{\rho p}{h}\sum_{i=2}^{N} q_{ii}\rho^2.$$

If $h \simeq \rho$, all the terms of the above expression have the same order ρ^2 and rewriting them proves *(2)*. When $h \simeq \rho^\alpha$, $1 < \alpha < \frac{3}{2}$, we keep the lower order term and we get *(3)*. □

Interpretation in 2D

According to Theorem 4.1 the two-dimensional neighborhood filter acts as an evolution PDE with two terms. The first term is proportional to the second derivative of u in the direction $\xi = Du(\mathbf{x})^\perp/|Du(\mathbf{x})|$, which is tangent to the level line passing through $\mathbf{x}$. The second term is proportional to the second derivative of u in the direction $\eta = Du(\mathbf{x})/|Du(\mathbf{x})|$ which is orthogonal to the level line passing through $\mathbf{x}$. Like in the one dimensional case, the evolution equations $u_t = c_1u_{\xi\xi}$ and $u_t = c_2u_{\eta\eta}$ act as filtering or enhancing models depending on the signs of c_1 and c_2. Following the previous theorem, we can distinguish three cases, depending on the values of h and ρ.

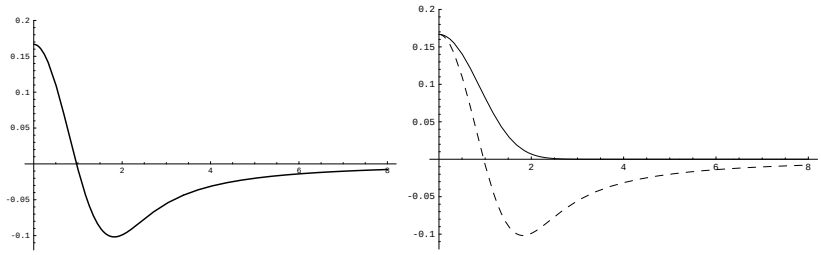

Figure 4: Weight functions of Theorems 3.1 and 4.1 when h and ρ have the same order. Left: Function f of Theorem 3.1. Right: Functions $\tilde{g}$ (continuous line) and $\tilde{f}$ (dashed line) of Theorem 4.1.

First, if h is much larger than ρ, both second derivatives are weighted by the same positive constant. Thus, the sum of the both terms is equivalent to the Laplacian of u, Δu, and we get back to Gaussian filtering.

Second, if h and ρ have the same order of magnitude, the neighborhood filter behaves as a filtering/enhancing algorithm. The coefficient of the diffusion in the tangential direction, $u_{\xi\xi}$, is given by $\tilde{g}(\frac{\rho}{h}|Du|)$. The function $\tilde{g}$ is positive and decreasing. Thus, there is always diffusion in that direction. The weight of the normal diffusion, $u_{\eta\eta}$, is given by $\tilde{f}(\frac{\rho}{h}|Du|)$. As the function $\tilde{f}$ takes positive and negative values (see figure 4), the filter behaves as a filtering/enhancing algorithm in the normal direction and depending on $|Du|$. If $\tilde{B}$ denotes the zero of $\tilde{f}$, then a filtering model is applied wherever $|Du| < \tilde{B}\frac{h}{\rho}$ and an enhancing strategy wherever $|Du| > \tilde{B}\frac{h}{\rho}$. The intensity of the filtering in the tangent diffusion and the enhancing in the normal diffusion tend to zero when the gradient tends to infinity. Thus, points with a very large gradient are not altered.

Finally, if ρ is much larger than h, the value $\frac{\rho}{h}$ tends to infinity and then the filtering magnitude $\tilde{g}(\frac{\rho}{h}|Du|)$ tends to zero. Thus, the original image is hardly altered. Let us mention that similar calculations were performed in a particular case for the neighborhood median filter by Masnou [16].

We observe that when ρ and h have the same order, the neighborhood filter asymptotically behaves like a Perona-Malik model. Let us be more specific about this comparison. Taking $g(s) = \tilde{g}(s^{\frac{1}{2}})$ in the Perona-Malik equation (5), we obtain

$$u_t = \tilde{g}(|Du|)u_{\xi\xi} + \tilde{h}(|Du|)u_{\eta\eta}, \tag{9}$$

where $\tilde{h}(s) = \tilde{g}(s) + s\tilde{g}'(s)$. Thus, the Perona-Malik model and the neighborhood filter can be decomposed in the same way and with exactly the same weight in the tangent direction. Then the function $\tilde{h}$ has the same behavior as $\tilde{f}$ (Theorem 4.1), as can be observed in Figure 5. Thus, in this case, a neighborhood filter has the same qualitative behavior as a Perona-Malik model, even if we cannot rewrite it exactly as such.

Figure 8 displays a comparison of the neighborhood filter and the Perona-Malik model. We display a natural image and the filtered images by both models. These solutions have a similar visual quality and tend to display flat zones and artificial contours inside the smooth regions. Figure 9 corroborates this visual impression. We display the level lines of both filtered solutions. As expected from the above consistency theorems, for both models the level lines of the original image tend to concentrate, thus creating large flat zones separated by edges. The solutions are very close, up to the obvious very different implementations. The neighborhood filter is implemented exactly as in its definition and the Perona-Malik model by the explicit difference scheme proposed in the original paper.

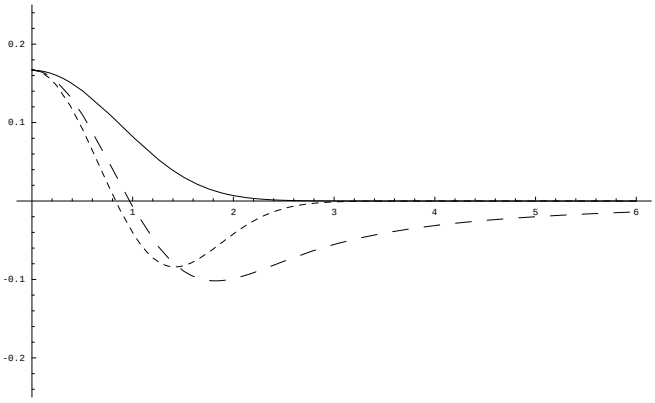

Figure 5: Weight comparison of the neighborhood filter and the Perona Malik equation. Magnitude of the tangent diffusion (continuous line, identical for both models) and normal diffusion (dashed line – –) of Theorem 4.1. Magnitude of the tangent diffusion (continuous line) and normal diffusion (dashed line - - -) of the Perona-Malik model (9). Both models show nearly the same behavior.

5 A regression correction of the neighborhood filter

In the previous sections we have shown the enhancing character of the neighborhood filter. We have seen that the neighborhood filter, like the Perona-Malik model, can create large flat zones and spurious contours inside smooth regions. This effect depends upon a gradient threshold which is hard to fix in such a way as to always separate the visually smooth regions from edge regions. In order to avoid this undesirable effect, let us analyze in more detail what happens with the neighborhood filter in the one-dimensional case.

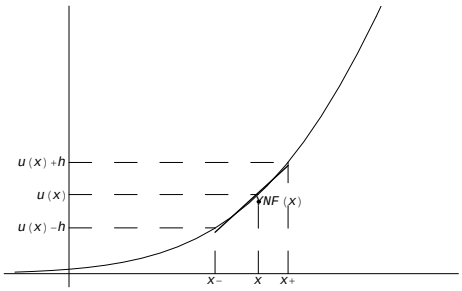

Figure 6: Illustration of the shock effect of the YNF on the convex of a signal. The number of points y satisfying $u(x) - h < u(y) \leq u(x)$ is larger than the number satisfying $u(x) \leq u(y) < u(x) + h$. Thus, the average value $YNF(x)$ is smaller than $u(x)$, enhancing that part of the signal. The regression line of u inside (x_-, x_+) better approximates the signal at x.

Figure 6 shows a simple illustration of this effect. For each x in the convex part of the signal, the filtered value is the average of the points y such that $u(x) - h < u(y) < u(x) + h$ for a certain threshold h. As it is illustrated in the figure, the number of points satisfying $u(x) - h < u(y) \leq u(x)$ is larger than the number of points satisfying $u(x) \leq u(y) < u(x) + h$. Thus, the average value $YNF(x)$ is smaller than $u(x)$, enhancing this part of the signal. A similar argument can be applied in the concave parts of the signal, dealing to the same enhancing effect. Therefore, shocks will be created inside smooth zones where concave and convex parts meet. Figure 6 also shows how the mean is not

a good estimate of $u(x)$ in this case. In the same figure, we display the regression line approximating u inside $(u^{-1}(u(x)-h), u^{-1}(u(x)+h))$. We see how the value of the regression line at x better approximates the signal. In the sequel, we propose to correct the neighborhood filter with this better estimate.

In the general case, this linear regression strategy amounts to finding for every point $\mathbf{x}$ the hyperplane locally approximating u in the following sense,

$$\min_{a_0,a_1,\ldots,a_N} \int_{B_\rho(\mathbf{x})} w(\mathbf{x},\mathbf{y})(u(\mathbf{y}) - a_N y_N - \ldots - a_1 y_1 - a_0)^2 \, d\mathbf{y}, \qquad w(\mathbf{x},\mathbf{y}) = e^{-\frac{|u(\mathbf{y})-u(\mathbf{x})|^2}{h^2}} \tag{10}$$

and then replacing $u(\mathbf{x})$ by the filtered value $a_N x_N + \ldots + a_1 x_1 + a_0$. The weights used to define the minimization problem are the same as the ones used by the neighborhood filter. Thus, the points with a grey level value close to $u(x)$ will have a larger influence in the minimization process than those with a further grey level value. We denote the above linear regression correction by $LYNF_{h,\rho}$. Taking $a_1 = \ldots = a_N = 0$ and then approximating u by a constant function, the minimization (10) goes back to the neighborhood filter.

In a more general framework, we can locally approximate image u by any polynomial. For any $n > 1$, we can find the polynomial p_n of degree n minimizing the following error,

$$\min_{p_n} \int_{B_\rho(\mathbf{x})} w(x,y)(u(y) - p_n(y_1,\ldots,y_N)^2 dy, \qquad w(x,y) = e^{-\frac{|u(y)-u(x)|^2}{h^2}} \tag{11}$$

and define the filtered value at $\mathbf{x}$ as $p_n(x_1,\ldots,x_N)$. We denote the above polynomial regression correction as $LYNF_{n,h,\rho}$.

A similar scheme has been statistically studied in [22]. The authors propose an iterative procedure that describes for every point the largest possible neighborhood in which the initial data can be well approximated by a parametric function.

Another similar strategy is the interpolation by ENO schemes [8]. The goal of ENO interpolation is to obtain a better adapted prediction near the singularities of the data. For each point it selects different stencils of fixed size M, and for each stencil reconstructs the associated interpolation polynomial of degree M. Then the *least oscillatory* polynomial is selected by some prescribed numerical criterion. The selected stencils tend to escape from large gradients and discontinuities.

The regression strategy also tends to select the right points in order to approximate the function. Instead of choosing a certain interval, all the points are used in the polynomial reconstruction, but weighted by the grey level differences.

5.1 Linear regression

As in the previous sections, let us analyze the asymptotic behavior of the linear regression correction. We compute the asymptotic expansion of the filter when $0 < \alpha \leq 1$. We showed that when $\alpha > 1$ the signal is hardly modified.

For the sake of completeness, we first compute the asymptotic expansion in the one dimensional case. We omit the proof of this result since it can be easily obtained from the next proof in the N-dimensional case.

Theorem 5.1 *Suppose $u \in C^2(I)$, and let $\rho, h, \alpha > 0$ such that $\rho, h \to 0$ and $h = O(\rho^\alpha)$. Let $\tilde{f}$ be the continuous function defined as $\tilde{f}(0) = \frac{1}{6}$,*

$$\tilde{f}(t) = \frac{1}{4t^2}\left(1 - \frac{2t\, e^{-t^2}}{E(t)}\right),$$

for $t \neq 0$, where $E(t) = 2\int_0^t e^{-s^2}\, ds$. Then, for $x \in \mathbb{R}$,

1. *If* $\alpha < 1$, $\quad LYNF_{h,\rho}u(x) - u(x) \simeq \frac{u''(x)}{6}\rho^2$.

2. *If* $\alpha = 1$, $\quad YNF_{h,\rho}u(x) - u(x) \simeq \tilde{f}(\frac{\rho}{h}\,|u'(x)|)\,u''(x)\,\rho^2$.

Theorem 5.1 shows that the $LYNF_{h,\rho}$ filter lets the signal evolve proportionally to its second derivative, as the neighborhood filter does. When h is larger than ρ the filter is equivalent to the original neighborhood filter and the signal is filtered by a heat equation. When ρ and h have the same order the sign and magnitude of the filtering process is given by $\tilde{f}(\frac{\rho}{h}\,|u'(x)|)$ (see Figure 7). This function is positive and quickly decreases to zero. Thus, the signal is filtered by a heat equation of decreasing magnitude and is not altered wherever the derivative is very large.

The same asymptotic expansion can be computed in the N-dimensional case. Let $e_1 = Du(\mathbf{x})/|Du(\mathbf{x})|$ and take $e_2, \ldots, e_N$ as an orthonormal basis of the hyperplane orthogonal to Du, $Du^{\perp}$. Then the set $\{e_1, e_2, \ldots, e_N\}$ forms an orthonormal basis of $\mathbb{R}^N$.

Theorem 5.2 *Suppose* $u \in C^2(\Omega)$, *and let* $\rho, h, \alpha > 0$ *such that* $\rho, h \to 0$ *and* $h = O(\rho^{\alpha})$. *Let* $\tilde{f}$ *be the continuous function defined as* $\tilde{f}(0) = \frac{1}{6}$,

$$\tilde{f}(t) = \frac{1}{4t^2}\left(1 - \frac{2t\,e^{-t^2}}{E(t)}\right),$$

for $t \neq 0$, *where* $E(t) = 2\int_0^t e^{-s^2}\,ds$. *Then, for* $\boldsymbol{x} \in \Omega$,

1. *If* $\alpha < 1$,

$$LYNF_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq \frac{\triangle u(\boldsymbol{x})}{6}\rho^2.$$

2. *If* $\alpha = 1$,

$$LYNF_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq \left[\tilde{f}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)\;D^2u(e_1,e_1)(\boldsymbol{x}) + \frac{1}{6}\sum_{i=2}^{N} D^2u(e_i,e_i)\right]\rho^2$$

Proof: The solution of the minimization process (10) leads to the resolution of the linear system $Ax = b$,

$$A = \begin{pmatrix} a(2,0,\ldots,0) & a(1,1,\ldots,0) & \ldots & a(1,0,\ldots,1) & a(1,0,\ldots,0) \\ a(1,1,\ldots,0) & a(0,2,\ldots,0) & \ldots & a(0,1,\ldots,1) & a(0,1,\ldots,0) \\ \vdots & & \ddots & & \vdots \\ a(1,0,\ldots,1) & a(0,1,\ldots,1) & \ldots & a(0,\ldots,0,2) & a(0,0,\ldots,1) \\ a(1,0,\ldots,0) & a(0,1,\ldots,0) & \ldots & a(0,0,\ldots,1) & a(0,0,\ldots,0) \end{pmatrix} \quad b = \begin{pmatrix} b(1,0,\ldots,0) \\ b(0,1,\ldots,0) \\ \vdots \\ b(0,0,\ldots,1) \\ b(0,0,\ldots,0) \end{pmatrix}$$

where

$$a(\alpha_1,\ldots,\alpha_N) = \int_{B_\rho(\mathbf{x})} t_1^{\alpha_1}\ldots t_N^{\alpha_N}\,w(t_1,\ldots,t_N)\,dt_1\ldots dt_N,$$

$$b(\alpha_1,\ldots,\alpha_N) = \int_{B_\rho(\mathbf{x})} t_1^{\alpha_1}\ldots t_N^{\alpha_N}\,u(t_1,\ldots,t_N)\,w(t_1,\ldots,t_N)\,dt_1\ldots dt_N,$$

$\mathbf{x} = (x_1,\ldots,x_N)$ and $w(t_1,\ldots,t_N) = e^{-\frac{|u(t_1,\ldots,t_N)-u(x_1,\ldots,x_N)|^2}{h^2}}$.

We can suppose, without loss of generality, that $\mathbf{x} = 0$. In this case, $LYNF_{h,\rho}u(0) = a_0$ can be written as $\det A'/\det A$, where A is defined above and A' is obtained by replacing the last column of A by b. Now, by taking the function $\tilde{u}(\mathbf{t}) = u(\mathbf{t}) - u(0)$ we can rewrite

$$b(\alpha_1,\ldots,\alpha_N) = \tilde{b}(\alpha_1,\ldots,\alpha_N) + u(0)\,a(\alpha_1,\ldots,\alpha_N),$$

and $\det A'$ can be decomposed as $\det \tilde{A} + u(0) \det A$. Then the difference between the filtered and original images is

$$LYNF_{h,\rho}u(0) - u(0) = \frac{\det \tilde{A} + u(0) \det A}{\det A} - u(0) = \frac{\det \tilde{A}}{\det A}.$$

We take the Taylor expansion of $\mathrm{u}(\mathbf{t})$ for $\mathbf{t} \in B_\rho(0)$,

$$u(\mathbf{t}) = u(0) + pt_1 + \sum_{i,j \in \{1,\dots,N\}} q_{ij} t_i t_j + O(|\mathbf{t}|^3),$$

where $\mathbf{t} = (t_1, \dots, t_N)$, $p = |Du(0)|$ and if $p > 0$,

$$q_{ii} = \frac{1}{2} D^2 u(e_i, e_i), \quad q_{ij} = D^2 u(e_i, e_j) \quad \text{if} \quad i \neq j.$$

When $\alpha < 1$, we apply the usual Taylor expansion of the exponential function. The terms of lower order of matrices A and $\tilde{A}$ are in their diagonal and the quotient can be approximated by the lower terms of

$$\frac{\tilde{b}(0, 0, \dots, 0)}{a(0, 0, \dots, 0)}.$$

Therefore, the analysis of the difference reduces to the computation of the two terms,

$$a(0, \dots, 0) \simeq \int_{B_\rho(0)} dt_1 \dots dt_N \simeq 2^N \rho^N$$

$$\tilde{b}(0, \dots, 0) \simeq \sum_{i=1}^{N} q_{ii} \int_{B_\rho(0)} t_i^2 dt_1 \dots dt_N = \frac{2^N \rho^{N+2}}{6} \Delta u$$

This proves *(1)*.

When $\alpha = 1$, we cannot apply the above expansion because $\frac{\rho^2}{h^2}$ does not tend to zero. However, $\frac{\rho^3}{h^2} \to 0$, and we can decompose the exponential as

$$e^{-\frac{|u(\mathbf{t}) - u(0)|^2}{h^2}} = e^{\frac{-p^2 t_1^2}{h^2}} \left(1 - \frac{2pt_1}{h^2} \sum_{i,j} q_{ij} t_i t_j + O(\frac{|t|^4}{h^2}) \right).$$

The lower order terms of the matrices A and $\tilde{A}$ are the diagonal elements, $a(1, 0, \dots, 0)$ and $\tilde{b}(1, 0, \dots, 0)$. Then, the lower order terms of the quotient are given by

$$\begin{aligned} \frac{\det \tilde{A}}{\det A} &\simeq \frac{\tilde{b}(0, \dots, 0) \prod_{i=1}^{N} a(0, \dots, \overset{(i)}{2}, \dots, 0) - \tilde{b}(1, 0, \dots, 0) a(1, 0, \dots, 0) \prod_{i=2}^{N} a(0, \dots, \overset{(i)}{2}, \dots, 0)}{a(0, \dots, 0) \prod_{i=1}^{N} a(0, \dots, \overset{(i)}{2}, \dots, 0)} \\ &= \frac{\tilde{b}(0, \dots, 0) a(2, 0, \dots, 0) - \tilde{b}(1, 0, \dots, 0) a(1, 0, \dots, 0)}{a(0, \dots, 0) a(2, 0, \dots, 0)}. \end{aligned} \tag{12}$$

Therefore, the analysis of the difference reduces to the computation of the terms,

$$a(0, \dots, 0) \simeq \int_{B_\rho(0)} e^{\frac{-p^2 t_1^2}{h^2}} dt_1 \dots dt_N$$

$$a(1,0,\ldots,0) \simeq -\frac{2p}{h^2}\sum_{i=1}^{N} q_{ii}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2\, t_i^2\, dt_1\ldots dt_N$$

$$a(2,0,\ldots,0) \simeq \int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2\, dt_1\ldots dt_N$$

$$\tilde{b}(0,\ldots,0) \simeq \sum_{i=1}^{N} q_{ii}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_i^2\, dt_1\ldots dt_N - \frac{2p^2}{h^2}\sum_{i=1}^{N} q_{ii}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2 t_i^2\, dt_1\ldots dt_N$$

$$\tilde{b}(1,0,\ldots,0) \simeq p\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2\, dt_1\ldots dt_N.$$

Now, replacing the terms in (12) by the previous estimates we get

$$\begin{aligned} \mathit{LYNF}_{h,\rho}u(0) - u(0) \quad &\simeq \quad \frac{q_{11}}{\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}}\, dt_1\ldots dt_N}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_1^2\, dt_1\ldots dt_N \\ &+ \quad \sum_{i=2}^{N} \frac{q_{ii}}{\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}}\, dt_1\ldots dt_N}\int_{B_\rho(0)} e^{\frac{-p^2t_1^2}{h^2}} t_i^2\, dt_1\ldots dt_N. \end{aligned}$$

Computing the previous integrals and taking into account that $O(\rho) = O(h)$ we prove *(2)*. □

Interpretation in 2D

According to the previous theorem, the filter can be written as the sum of two diffusion terms in the direction of ξ and η. When h is much larger than ρ the linear regression correction is equivalent to the heat equation like the original neighborhood filter. When ρ and h have the same order, the behavior of the linear regression algorithm is very different from the original neighborhood filter. The function weighting the tangent diffusion is a positive constant. The function weighting the normal diffusion is positive and decreasing (see figure 7), and therefore there is no enhancing effect. The algorithm combines the tangent and normal diffusion wherever the gradient is small. Wherever the gradient is larger the normal diffusion is canceled and the image is filtered only in its tangent direction. This subjacent PDE was already proposed as a diffusion equation in [13]. This diffusion makes the level lines evolve proportionally to their curvature. In the Perona-Malik model the diffusion is stopped near the edges. In this case, the edges are filtered by a mean curvature motion.

It may be asked whether the modified neighborhood filter still preserves signal discontinuities. The answer is yes. It is easily checked that for small enough h, all piecewise affine functions with smooth jump curves are steady. Thus, the behavior is the same as for the classical neighborhood filter. Our asymptotic analysis is of course not valid for such functions, but only for smooth functions.

As a numerical scheme the linear regression neighborhood filter allows the implementation of a mean curvature motion without the computation of gradients and orientations. When the gradient is small the linear regression filter naturally behaves like the heat equation. This effect is introduced on typical schemes implementing the mean curvature motion. In flat zones the gradient is not well defined and some kind of isotropic diffusion must be applied. Therefore, the linear regression neighborhood filter naturally extends the mean curvature motion and yields a stable numerical scheme for its computation, independent of gradient orientations.

Figure 8 displays an experiment comparing the $\mathit{LYNF}_{h,\rho}$ with the neighborhood filter and the Perona-Malik equation. The linear correction does not create any contour or flat zone inside the smooth regions. Figure 9 displays the level lines of the previous experiment. The level lines of the $\mathit{LYNF}_{h,\rho}$ are filtered by a mean curvature motion and they do not get grouped creating flat zones.

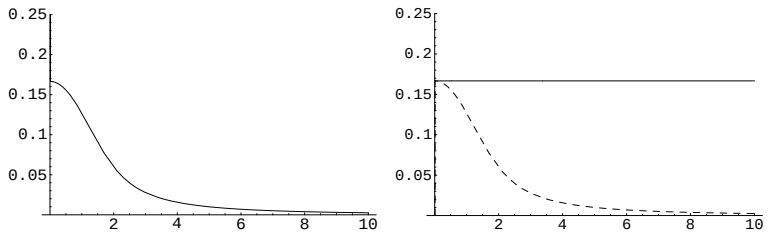

Figure 7: Weighting functions of Theorem 5.1 and 5.2. Left: Function $\tilde{f}$ of Theorem 5.1. Right: Constant function 1/6 (continuous line) and function $\tilde{f}$ (dashed line) of Theorem 5.2.

In the N-dimensional case, we can choose $\{e_2, \ldots, e_N\}$ to be a basis of eigenvectors of $D^2u(\mathbf{x})$ restricted to $Du(\mathbf{x})^\perp$. In this case the diffusion terms are equal to the principal curvatures of the level surface passing through $\mathbf{x}$, and their sum gives the mean curvature operator. Therefore when the gradient is above a certain threshold the linear regression neighborhood filter is equivalent to the N-dimensional mean curvature motion.

Figure 10 shows a benefit of the linear regression correction in the three-dimensional case. We apply the neighborhood filter to the 3D image $u(x, y, z) = x^2 + y^2 + z^2 - r^2$. The zero level set of u is the boundary of the sphere of radius r. We should expect this level set to evolve isotropically and to preserve its sphere form, which is not the case when applying the neighborhood filter. This effect is avoided by the linear regression which makes the zero level set evolve uniformly and preserve the isotropy of the initial data.

5.2 The action of the regression filter on image sequences

Classical movie filtering techniques usually compute an optical flow estimate of the sequence as a previous step. The optical flow associates to each point a vector representing its optical velocity, $v = (v_x, v_y)$ (we denote $\mathbf{v} = (v_x, v_y, 1)$). If Δt is the time interval between two consecutive frames, $\mathbf{x} + \mathbf{v}(\mathbf{x})\Delta t$ denotes the point shifted by $\mathbf{v}(\mathbf{x})$ in the next frame. The optical flow estimation gives the theoretical trajectories of each pixel along the sequence. Thus, denoising a pixel should be performed by averaging its computed trajectory.

The classical definition of the optical flow involves the Lambertian assumption that physical points maintain the same grey level value when the camera moves,

$$u(\mathbf{x} + \mathbf{v}(\mathbf{x})\Delta t) \simeq u(\mathbf{x}),$$

which is equivalent to saying that $\mathbf{v}(\mathbf{x}) \cdot Du(\mathbf{x}) = 0$. Therefore, the conservation assumption only tells us that the velocity vector is contained in the plane $Du^\perp$. Some additional criteria must be added in order to achieve the uniqueness of the trajectories. Theorem 5.2 shows that when the gradient is above a certain threshold the linear regression neighborhood filter performs a diffusion in the plane $Du^\perp$. Thus, instead of choosing pixels of the trajectory for the averaging process, which would require a solution to the ambiguity of trajectories, the algorithm makes an average of all the points with a similar grey level value.

The enhancing effect of the neighborhood filter is aggravated by the presence of noise. The irregularities of the edges due to the noise are enhanced and lead to very irregular and oscillatory contours. This effect is even more annoying when dealing with films. The irregularities of a given edge are different from frame to frame, leading to a false motion impression when the sequence is played. In Figure 12 we display two consecutive frames of a noisy image sequence. We display the same frames once the sequence has been filtered by the neighborhood filter and its linear correction. On comparing the single images we observe that the edges of the filtered sequence by the neighborhood filter are

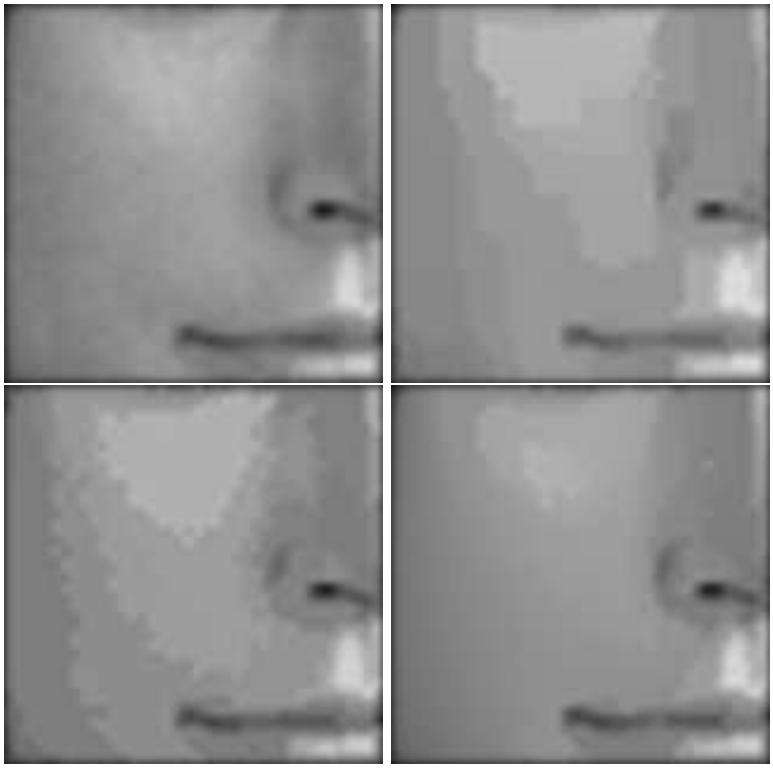

Figure 8: Comparison experiment. Top left: original image. Top right: Perona-Malik filtered image. Bottom left: filtered image by the neighborhood filter. Bottom right: filtered image by the linear regression neighborhood filter. The neighborhood filter experiments are performed by iterating the discrete version of definitions (1) and (10). Both the neighborhood filter and its linear regression correction have been applied with the same value of h and ρ. The displayed images have been attained within the same number of iterations. The Perona-Malik equation is implemented by the explicit difference scheme proposed in the original paper. The Perona-Malik model and the neighborhood filter create artificial contours and flat zones. This effect is almost completely avoided by the linear regression neighborhood filter.

much more irregular than the original ones. The linear regression correction avoids this effect and regularizes the edges. Finally, we display for each filtered sequence the difference between the two consecutive frames for each filtered sequence. The edges are less noticeable in the frame difference of the linear regression correction, that is, the oscillations from frame to frame are reduced.

6 The vector valued case

Let u be a vector valued function defined on a bounded domain $\Omega \subset \mathbb{R}^2$, $u : \Omega \to \mathbb{R}^n$. The vector neighborhood filter can be written as

$$YNF_{h,\rho}u(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{B_\rho(\mathbf{x})} u(\mathbf{y}) e^{-\frac{||u(\mathbf{y})-u(\mathbf{x})||^2}{h^2}} d\mathbf{y}, \tag{13}$$

where $||u(\mathbf{y}) - u(\mathbf{x})||^2$ is now the Euclidean vector norm and each component function u_i is filtered with the same weight distribution. The linear regression correction is defined as in the scalar case,

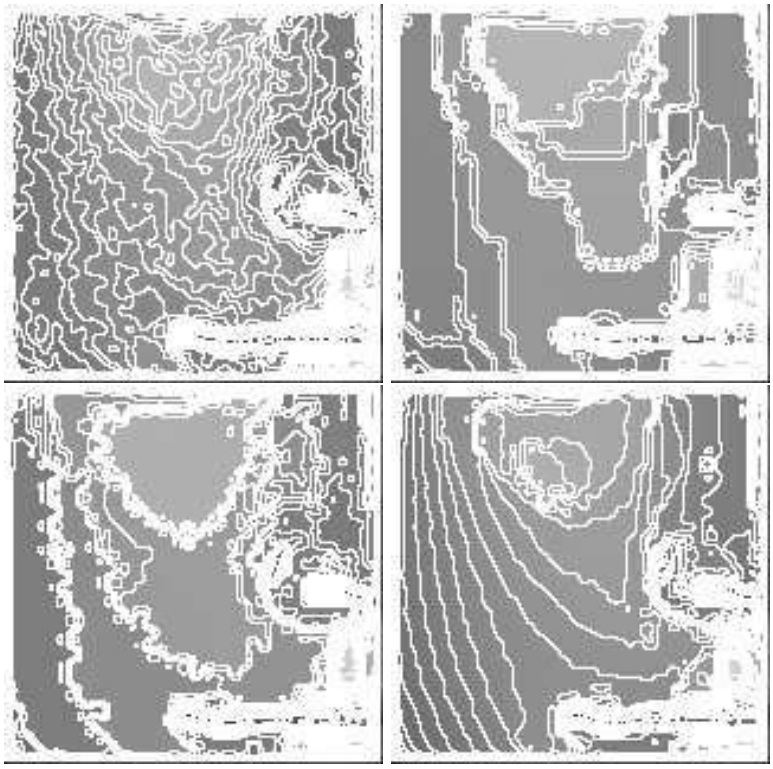

Figure 9: Level lines of the images in figure 8. By the Perona-Malik filter and the neighborhood filter the level lines tend to group, creating flat zones. The regression correction filters the level lines by a curvature motion without creating any flat zone.

and each component is locally approximated by a plane with the same weight distribution.

In order to compute the asymptotic expansion of the linear regression filter we must fix a coordinate system for $\mathbb{R}^2$. In the scalar case we used the reference system given by the gradient of the image at $\mathbf{x}$ and its orthogonal direction. In addition, this reference allows us to relate the obtained diffusion to the evolution of the level lines of the image and the mean curvature motion. Now, we cannot use the same reference and we need to define a new one. By analogy with the scalar case we choose the directions of minimum and maximum variation of the vector function.

Definition 6.1 *We define the normal direction η and the tangent direction ξ as the vectors that respectively maximize and minimize the following variation*

$$\sum_{i=1}^{n} \| \frac{\partial u_i}{\partial v}(\boldsymbol{x}) \|^2$$

under the constraint $\|v\| = 1$.

It is easily seen that this constrained optimization leads to the computation of the eigenvectors of the matrix

$$A = \begin{pmatrix} \| \frac{\partial u}{\partial x} \|^2 & \langle \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \rangle \\ \langle \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \rangle & \| \frac{\partial u}{\partial y} \|^2 \end{pmatrix},$$

Figure 10: Evolution of a 3D sphere by the linear regression neighborhood filter. Left: Zero level set of $u(x,y,z) = x^2+y^2+z^2-r^2$ (sphere of radius r). Middle: Zero level set of u after the iteration of the neighborhood filter. Right: Zero level set of u after the iteration of the linear regression neighborhood filter. The linear regression makes the zero level set evolve uniformly and preserves the isotropic property of the initial data, which is not the case of the neighborhood filter.

where $\frac{\partial u}{\partial x} = \left(\frac{\partial u_1}{\partial x}, \ldots, \frac{\partial u_n}{\partial x}\right)$ and $\frac{\partial u}{\partial y} = \left(\frac{\partial u_1}{\partial y}, \ldots, \frac{\partial u_n}{\partial y}\right)$. The two positive eigenvalues of A, λ_+ and λ_-, are the maximum and the minimum of the vector norm associated to A and the maximum and the minimum variations as defined in Definition 6.1. The corresponding eigenvectors are orthogonal leading to the above defined normal and tangent directions. This orthonormal system was first proposed for vector valued image analysis in [5]. Many PDE equations have been proposed for color image filtering using this system. We note the Coherence Enhancing Diffusion [27], the Beltrami Flow [11] and an extension of the mean curvature motion [21].

Theorem 6.1 *Suppose $u \in C^2(\Omega, \mathbb{R}^n)$, and let $\rho, h, \alpha > 0$ such that $\rho, h \to 0$ and $h = O(\rho^\alpha)$. Let $\tilde{f}$ be the continuous function defined as $\tilde{f}(0) = \frac{1}{6}$,*

$$\tilde{f}(t) = \frac{1}{4t^2}\left(1 - \frac{2t\, e^{-t^2}}{E(t)}\right),$$

for $t \neq 0$, where $E(t) = 2\int_0^t e^{-s^2}\, ds$. Then, for $\boldsymbol{x} \in \Omega$,

1. *If $\alpha < 1$,*

$$\mathit{LYNF}_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq \frac{\triangle u(\boldsymbol{x})}{6}\rho^2.$$

2. *If $\alpha = 1$,*

$$\mathit{LYNF}_{h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq \left[\tilde{f}(\frac{\rho}{h}\,\|\frac{\partial u}{\partial \xi}(\boldsymbol{x})\|)\; D^2u(\xi,\xi)(\boldsymbol{x}) + \tilde{f}(\frac{\rho}{h}\|\frac{\partial u}{\partial \eta}(\boldsymbol{x})\|)D^2u(\eta,\eta)(\boldsymbol{x})\right]\rho^2$$

where $\Delta u(\boldsymbol{x}) = (\Delta u_i(\boldsymbol{x}))_{1\leq i\leq n}$ and $D^2u(v,v)(\boldsymbol{x}) = \left(D^2u_i(v,v)(\boldsymbol{x})\right)_{1\leq i\leq n}$ for $v \in \{\eta, \xi\}$.

Proof: Let u_0 denote an arbitrary component of u and let us suppose without loss of generality that $\mathbf{x} = 0$. In this case, the same argument of the proof of Theorem 5.2 shows that

$$\mathit{LYNF}_{h,\rho}u_0(0) - u(0) = \frac{\det \tilde{A}}{\det A},$$

where

$$A = \begin{pmatrix} a(2,0) & a(1,1) & a(1,0) \\ a(1,1) & a(0,2) & a(0,1) \\ a(1,0) & a(0,1) & a(0,0) \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} a(2,0) & a(1,1) & \tilde{b}(1,0) \\ a(1,1) & a(0,2) & \tilde{b}(0,1) \\ a(1,0) & a(0,1) & \tilde{b}(0,0) \end{pmatrix}$$

and

$$a(\alpha_1,\alpha_2)=\int_{B_\rho(0)} t_1^{\alpha_1}t_2^{\alpha_2}w(t_1,t_2)\,dt_1dt_2,$$

$$\tilde{b}(\alpha_1,\alpha_2)=\int_{B_\rho(0)} t_1^{\alpha_1}t_2^{\alpha_2}w(t_1,t_2)\,[u_0(t_1,t_2)-u_0(0)]\,dt_1dt_2.$$

The weight function $w(t_1,t_2)$ depends on the differences $|u_i(t_1,t_2)-u_i(0)|$, for $i=1,\dots,n$,

$$w(t_1,t_2)=e^{-\frac{1}{h^2}\sum_{i=1}^n |u_i(t_1,t_2)-u_i(0,0)|}.$$

We take the Taylor expansion of $u_0(\mathbf{t})$ and $u_i(\mathbf{t})$ for $\mathbf{t}\in B_\rho(0)$,

$$u_i(\mathbf{t})=u_i(0)+p_{i_\xi}t_1+p_{i_\eta}t_2+q_{i_{\xi\xi}}t_1^2+q_{i_{\eta\eta}}t_2^2+q_{i_{\xi\eta}}t_1t_2+O(|\mathbf{t}|^3),$$

where $\mathbf{t}=(t_1,t_2)$, $p_{i_\xi}=\frac{\partial u_i}{\partial\xi}$, $p_{i_\eta}=\frac{\partial u_i}{\partial\eta}$ and

$$q_{i_{\xi\xi}}=\frac{1}{2}D^2u_i(\xi,\xi),\quad q_{i_{\eta\eta}}=\frac{1}{2}D^2u_i(\eta,\eta)\quad\text{and}\quad q_{i_{\xi\eta}}=D^2u_i(\xi,\eta).$$

When $\alpha<1$, we apply the usual Taylor expansion of the exponential function. The lower order terms of matrices A and $\tilde{A}$ are in their diagonal and the quotient can be approximated by the lower terms of $\tilde{b}(0,0)/a(0,0)$. Therefore, the analysis of the difference reduces to the computation of the two terms,

$$a(0,0)\simeq\int_{B_\rho(0)}dt_1dt_2\simeq 4\rho^2$$

$$\tilde{b}(0,0)\simeq\int_{B_\rho(0)}(q_{0_{\xi\xi}}t_1^2+q_{0_{\eta\eta}}t_2^2)\,dt_1dt_2=\frac{4\Delta u_0}{6}\rho^4$$

This proves *(1)*.

When $\alpha=1$, we cannot apply the above expansion and we decompose the weight function as

$$w(t_1,t_2)\simeq e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\left(1-\frac{1}{h^2}(c_{30}t_1^3+c_{21}t_1^2t_2+c_{12}t_1t_2^2+c_{03}t_2^3)\right),$$

where

$$a=\|\frac{\partial u}{\partial\xi}\|,\quad b=\|\frac{\partial u}{\partial\eta}\|,\quad c_{30}=2\langle D^2u(\xi,\xi),\frac{\partial u}{\partial\xi}\rangle,\quad c_{03}=2\langle D^2u(\eta,\eta),\frac{\partial u}{\partial\eta}\rangle,$$

$$c_{21}=2\langle D^2u(\xi,\eta),\frac{\partial u}{\partial\xi}\rangle+2\langle D^2u(\xi,\xi),\frac{\partial u}{\partial\eta}\rangle,\quad c_{12}=2\langle D^2u(\xi,\eta),\frac{\partial u}{\partial\eta}\rangle+2\langle D^2u(\eta,\eta),\frac{\partial u}{\partial\xi}\rangle.$$

We do not have a crossed term in the exponential function thanks to the orthogonality of $\frac{\partial u}{\partial\xi}$ and $\frac{\partial u}{\partial\eta}$. The lower order terms of the matrices A and $\tilde{A}$ are the diagonal elements, $a(1,0)$, $a(0,1)$, $\tilde{b}(1,0)$ and $\tilde{b}(0,1)$. Then, the lower order terms of the quotient are given by

$$\frac{\det\tilde{A}}{\det A}\simeq\frac{a(2,0)a(0,2)\tilde{b}(0,0)-a(1,0)a(0,2)\tilde{b}(1,0)-a(2,0)a(0,1)\tilde{b}(0,1)}{a(2,0)a(0,2)a(0,0)}.\tag{14}$$

Therefore, the analysis of the difference reduces to the computation of the terms,

$$a(0,0)\simeq\int_{B_\rho(0)}e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\,dt_1dt_2,$$

$$a(1,0) \simeq -\frac{1}{h^2}\int_{B_\rho(0)} (c_{30}t_1^4 + c_{12}t_1^2t_2^2)\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2,$$

$$a(0,1) \simeq -\frac{1}{h^2}\int_{B_\rho(0)} (c_{03}t_2^4 + c_{21}t_1^2t_2^2)\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2,$$

$$a(2,0) \simeq \int_{B_\rho(0)} t_1^2\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2, \quad a(0,2) \simeq \int_{B_\rho(0)} t_2^2\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2,$$

$$\tilde{b}(1,0) \simeq p_{0_\xi}\int_{B_\rho(0)} t_1^2\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2, \quad \tilde{b}(0,1) \simeq p_{0_\eta}\int_{B_\rho(0)} t_2^2\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2,$$

$$\begin{aligned}\tilde{b}(0,0) \quad \simeq \quad & \int_{B_\rho(0)} (q_{0_{\xi\xi}}t_1^2 + q_{0_{\eta\eta}}t_2^2) e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2 \\ & -\frac{1}{h^2}\int_{B_\rho(0)} (p_{0_\xi}c_{30}t_1^4 + p_{0_\eta}c_{03}t_2^4 + (p_{0_\xi}c_{12} + p_{0_\eta}c_{21})t_1^2t_2^2)\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2,\end{aligned}$$

Now, replacing the terms in (14) by the previous estimates we get

$$\frac{1}{\int_{B_\rho(0)} e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2}\int_{B_\rho(0)} (q_{0_{\xi\xi}}t_1^2 + q_{0_{\eta\eta}}t_2^2)\, e^{-\frac{1}{h^2}(a^2t_1^2+b^2t_2^2)}\, dt_1dt_2.$$

Computing the previous integrals and taking into account that $O(\rho) = O(h)$ we prove *(2)*. □

Interpretation
When h is much larger than ρ, the linear regression neighborhood filter is equivalent to the heat equation applied independently to each component. When h and ρ have the same order the subjacent PDE acts as an evolution equation with two terms. The first term is proportional to the second derivative of u in the tangent direction ξ. The second term is proportional to the second derivative of u in the normal direction η. The magnitude of each diffusion term depends on the variation in the respective direction, $\lambda_- = \|\frac{\partial u}{\partial \xi}(\mathbf{x})\|$ and $\lambda_+ = \|\frac{\partial u}{\partial \eta}(\mathbf{x})\|$. The weighting function $\tilde{f}$ is positive and decreases to zero (see Figure 7). We can distinguish the following cases depending on the values of λ_+ and λ_-.

- If $\lambda_+ \simeq \lambda_- \simeq 0$ then there are very few variations of the vector image u around $\mathbf{x}$. In this case, the linear regression neighborhood filter behaves like a heat equation with maximum diffusion coefficient $\tilde{f}(0)$.
- If $\lambda_+ >> \lambda_-$ then there are strong variations of u around $\mathbf{x}$ and the point may be located on an edge. In this case the magnitude $\tilde{f}(\frac{\rho}{h}\lambda_+)$ tends to zero and there is no diffusion in the direction of maximal variation. If $\lambda_- >> 0$ then $\mathbf{x}$ may be placed on an edge with different orientations depending on each component and the magnitude of the filtering in both directions tends to zero, so that the image is hardly altered. If $\lambda_- \simeq 0$ then the edges have similar orientations in all the components and the image is filtered by a directional Laplacian in the direction of minimal variation.
- If $\lambda_+ \simeq \lambda_- >> 0$ then we may be located on a saddle point and in this case the image is hardly modified. When dealing with multi-valued images one can think of the complementarity of the different channels leading to the perception of a corner.

In the scalar case the theorem gives back the result studied in the previous sections. The normal and tangent directions are respectively the gradient direction and the level line direction. In this case, $\frac{\partial u}{\partial \xi}(\mathbf{x}) = 0$ and $\frac{\partial u}{\partial \eta}(\mathbf{x}) = |Du(\mathbf{x})|$ and we get back to

$$LYNF_{h,\rho}u(\mathbf{x}) - u(\mathbf{x}) \simeq \left[\frac{1}{6}D^2u(\xi,\xi)(\mathbf{x}) + \tilde{f}(\frac{\rho}{h}|Du(\mathbf{x})|)D^2u(\eta,\eta)(\mathbf{x})\right]\rho^2.$$

7 The Polynomial Regression

In this section, we extend the previous asymptotic analysis to the polynomial regression defined in (11). First, we compute the asymptotic expansion for the polynomials of degree 2. We only analyze the two dimensional case since we are mainly interested in its interpretation for image filtering.

Theorem 7.1 *Suppose $u \in C^2(\Omega)$, and let $\rho, h, \alpha > 0$ such that $\rho, h \to 0$ and $h = O(\rho^\alpha)$. Let*

$$\tilde{f}(t) = \frac{4\,t^2\,\left(-3+t^2\right) + 2\,e^{t^2}\,t\,\left(6-t^2+2\,t^4\right)\,E(t) - 3\,e^{2\,t^2}\,E(t)^2}{96\,t^4\,\left(2\,t^2 + e^{t^2}\,t\,\left(1+2\,t^2\right)\,E(t) - e^{2\,t^2}\,E(t)^2\right)},$$

and

$$\tilde{g}(t) = \frac{1}{36t^2}\left(1 - \frac{2\,t\,e^{-t^2}}{E(t)}\right)$$

for $t \neq 0$, where $E(t) = 2\int_0^t e^{-s^2}\,ds$. Then, for $\boldsymbol{x} \in \Omega$,

1. *If $\alpha < 1$,*

$$YNF_{2,h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq -\left[\frac{1}{280}u_{\eta\eta\eta\eta} + \frac{1}{54}u_{\eta\eta\xi\xi} + \frac{1}{280}u_{\xi\xi\xi\xi}\right]\rho^4$$

2. *If $\alpha = 1$,*

$$\begin{aligned} LYNF_{2,h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \quad \simeq \quad & -\left[\tilde{f}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\eta\eta} + \tilde{g}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\xi\xi} + \frac{1}{280}u_{\xi\xi\xi\xi}\right]\rho^4 \\ & + c\left[f_1(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\eta} + f_2(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\xi\xi}\right]u_{\eta\eta}\,\rho^4 \\ & + c\left[f_3(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\xi\xi\xi} + f_4(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\xi}\right]u_{\eta\xi}\,\rho^4 \\ & + c\left[f_5(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\eta} + f_6(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\xi\xi}\right]u_{\xi\xi}\,\rho^4 \end{aligned}$$

where c denotes ρ/h. The graphs of the functions $f_1, \ldots, f_6$ are plotted in figure 11.

The proof of the theorem follows the same scheme as previous results. However, the large number of terms and the complexity of the weighting functions make it impossible to write the proof on paper. The graphs of the functions involved in the theorem are plotted in Figure 11.

The polynomial approximation better adapts to the local image configuration than the linear approximation. The order obtained by the polynomial regression of degree 2 is ρ^4 while the neighborhood filter and its linear correction raised an order ρ^2.

Interpretation

When h is much larger than ρ the image is filtered by a nearly isotropic diffusion. The diffusion is written as a constant combination of $u_{\eta\eta\eta\eta}$, $u_{\xi\xi\xi\xi}$ and $u_{\eta\eta\xi\xi}$. To the best of our knowledge, there is

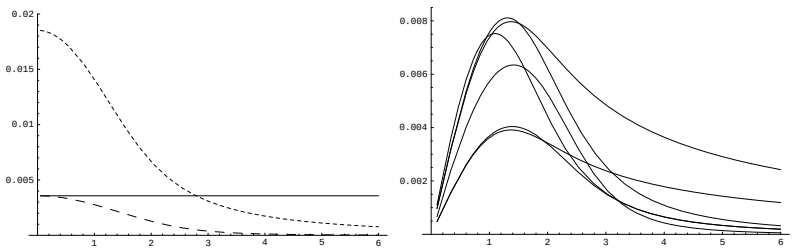

Figure 11: Weighting functions of Theorem 7.1. Left: Constant function 1/280 (continuous line), function $\tilde{f}$ (dashed line – –) and function $\tilde{g}$ (dashed line - - -). Right: Weighting functions $f_1, \ldots, f_6$.

no interpretation for this operator, even though it is quite similar to the bilaplacian operator. When h and ρ have the same order the polynomial regression acts as the combination of two terms. The first term is a fourth order term which becomes directional in the tangent direction ξ as the gradient increases (see Figure 11). The second term is formed by the product of derivatives of order 2 and 3. This term only acts when the gradient takes an intermediate value and cancels when it is very small or large (see Figure 11). When the gradient takes intermediate values the second terms acts, but we are not able to interpret this operator. This is why we computed the asymptotic expansion of the neighborhood filter with polynomial regression by polynomials of degree 3.

Theorem 7.2 *Suppose $u \in C^2(\Omega)$, and let $\rho, h, \alpha > 0$ such that $\rho, h \to 0$ and $h = O(\rho^\alpha)$. Let*

$$\tilde{f}(t) = \frac{4\,t^2\,\left(-3+t^2\right) + 2\,e^{t^2}\,t\,\left(6-t^2+2\,t^4\right)\,E(t) - 3\,e^{2\,t^2}\,E(t)^2}{96\,t^4\,\left(2\,t^2 + e^{t^2}\,t\,\left(1+2\,t^2\right)\,E(t) - e^{2\,t^2}\,E(t)^2\right)},$$

and

$$\tilde{g}(t) = \frac{1}{36t^2}\left(1 - \frac{2\,t\,e^{-t^2}}{E(t)}\right)$$

for $t \neq 0$, where $E(t) = 2\int_0^t e^{-s^2}\,ds$. Then, for $\boldsymbol{x} \in \Omega$,

1. *If $\alpha < 1$,*

$$YNF_{3,h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq -\left[\frac{1}{280}u_{\eta\eta\eta\eta} + \frac{1}{54}u_{\eta\eta\xi\xi} + \frac{1}{280}u_{\xi\xi\xi\xi}\right]\rho^4$$

2. *If $\alpha = 1$,*

$$LYNF_{3,h,\rho}u(\boldsymbol{x}) - u(\boldsymbol{x}) \simeq -\left[\tilde{f}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\eta\eta} + \tilde{g}(\frac{\rho}{h}\,|Du(\boldsymbol{x})|)u_{\eta\eta\xi\xi} + \frac{1}{280}u_{\xi\xi\xi\xi}\right]\rho^4$$

where c denotes ρ/h.

In the case $\alpha < 1$, we get the same approximation as in the polynomial case of degree 2. When $\alpha = 1$, by increasing the degree of the polynomial regression to 3, many terms of theorem 7.1 have been happily canceled and one obtains the same fourth order term. When the gradient is large, the subjacent PDE would again be $\frac{\partial u}{\partial t} = -u_{\xi\xi\xi\xi}$. This equation has not been studied, to the best of our knowledge.

8 Conclusion

Our first aim was to understand neighborhood filters thanks to their asymptotic PDE behavior. This led us to improve these filters by using the well-posed nature of the PDE as a cue. Conversely, this study introduced neighborhood filters as new kinds of numerical schemes for classical PDE's like the Perona-Malik or the mean curvature motion. The study led to two equations of interest,

$$\frac{\partial u}{\partial t} = -u_{\xi\xi\xi\xi}$$

and

$$\frac{\partial u}{\partial t} = -\left[\frac{1}{280}u_{\eta\eta\eta\eta} + \frac{1}{54}u_{\eta\eta\xi\xi} + \frac{1}{280}u_{\xi\xi\xi\xi}\right]$$

where ξ is the direction orthogonal to the gradient and η the direction of the gradient. We know of no existence theory for these equations. The two preceding theorems give them consistent numerical schemes.

References

[1] H. Amann, "A new approach to quasilinear parabolic problems", International Confererence on Differential Equations, 2005.

[2] F. Andreu, C. Ballester, V. Caselles and J. M. Mazï¿½, "Minimizing Total Variation Flow", Differential and Integral Equations, 14 (3), pp. 321-360, 2001.

[3] D. Barash, "A Fundamental Relationship between Bilateral Filtering, Adaptive Smoothing and the Nonlinear Diffusion Equation", IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (6), 2002.

[4] T.F. Chan, S. Esedoglu, F.E. Park, "Image Decomposition Combining Staircase Reduction and Texture Extraction", Preprint CAM-18, UCLA, 2005.

[5] S. Di Zenzo. "A note on the gradient of a multi-image". Computer Vision, Graphics, and Image Processing, vol. 33, pp. 116-125, 1986.

[6] S. Esedoglu, "An analysis of the Perona-Malik scheme", Communications on Pure and Applied Mathematics, vol. 54, pp. 1442-1487, 2001.

[7] F. Guichard and J.M Morel, "A Note on Two Classical Enhancement Filters and Their Associated PDE's", International Journal of Computer Vision, Volume 52 (2-3), pp. 153-160, 2003.

[8] A. Harten, B. Enquist, S. Osher and S. Chakravarthy, "Uniformly high order accurate essentially non-oscillatory schemes III", Journal of Computational Physics, vol. 71, pp. 231-303, 1987.

[9] S. Kichenassamy, "The Perona-Malik paradox", SIAM Journal Applied Mathematics, vol. 57 (2), pp. 1328-1342, 1997.

[10] B.B. Kimia, A. Tannenbaum, and S.W. Zucker, "On the evolution of curves via a function of curvature I the classical case", Journal of Mathematical Analysis and Applications, 163(2), pp. 438-458, 1992.

[11] R. Kimmel, R. Malladi, and N. Sochen, "Images as embedded maps and minimal surfaces: movies, color, texture, and volumetric medical images", International Journal of Computer Vision, vol. 39(2), pp. 111-129, 2000.

[12] S. Kindermann, S. Osher and P. Jones, "Deblurring and Denoising of Images by Nonlocal Functionals", UCLA Computational and Applied Mathematics Reports, 04-75, 2004.

[13] P. Kornprobst, "Contributions ï¿½ la Restauration d'Images et l'Analyse de Sï¿½uences: Approches Variationnelles et Solutions de Viscositï¿½. PhD thesis, Universitï¿½ de Nice-Sophia Antipolis, 1998.

[14] H. P. Kramer and J. B Bruckner. "Iterations of a non-linear transformation for enhancement of digital images". Pattern Recognition, 7, 1975.

[15] J.S. Lee "Digital image smoothing and the sigma filter", Computer Vision, Graphics, and Image Processing, vol. 24, pp. 255-269, 1983.

[16] S. Masnou, "Filtrage et désocclusion d'images par méthodes d'ensembles de niveau", PhD Dissertation, Université Paris-IX Dauphine, 1998.

[17] S. Osher and L. Rudin. "Feature oriented image enchancement using shock filters" SIAM J. Numerical Analysis, 27, pp. 919-940. 1990.

[18] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," IEEE Trans. Patt. Anal. Mach. Intell., 12, pp. 629-639, 1990.

[19] L. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms", Physica D, 60, pp. 259-268, 1992.

[20] P. Saint-Marc, J.S. Chen, and G. Medioni, "Adaptive Smoothing: A General Tool for Early Vision," IEEE Trans. Pattern Analysis and Machine Intelligence vol. 13 (6), pp. 514, 1991.

[21] G. Sapiro and D.L. Ringach, "Anisotropic diffusion of multivalued images with applications to color filtering", IEEE Transactions on Image Processing, vol. 5(11), pp. 1582-1585, 1996.

[22] J. Polzehl, V. Spokoiny, "Varying coefficient regression modeling", Preprint, Weierstrass Institute for Applied Analysis and Stochastics, 818, 2003.

[23] J.G.M. Schavemaker, M.J.T. Reinders, J.J. Gerbrands and E. Backer, "Image sharpening by morphological filtering", Pattern Recognition, 33, pp. 997-1012, 2000.

[24] J. Sethian "Curvature and the evolution of fronts", Comm. Math. Phys., 101, 1985.

[25] S.M. Smith and J.M. Brady, "Susan - a new approach to low level image processing," International Journal of Computer Vision, Volume 23 (1), pp. 45-78, 1997.

[26] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," Sixth International Conference on Computer Vision, pp. 839-46. 1998.

[27] J. Weickert, "Anisotropic Diffusion in Image Processing". Tuebner Stuttgart, 1998.

[28] L.P. Yaroslavsky, *Digital Picture Processing - An Introduction*, Springer Verlag, 1985.

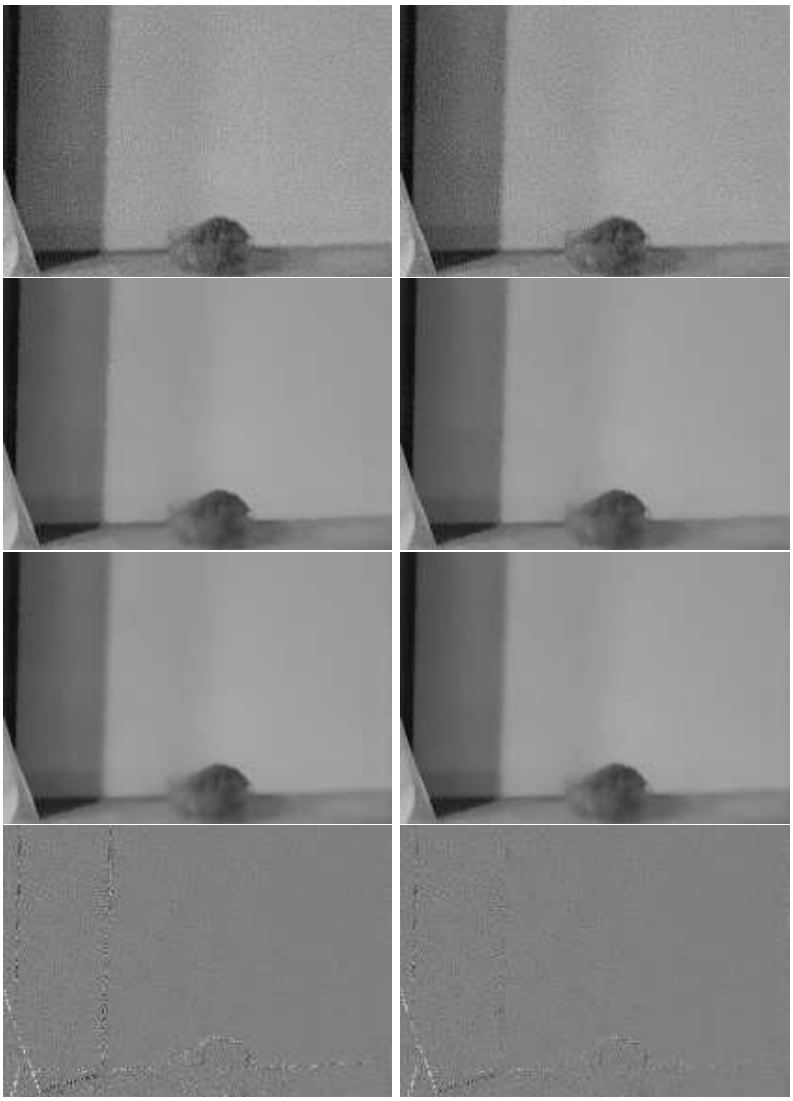

Figure 12: Experiment comparing the neighborhood filter and its linear correction on image sequences. From top to bottom: Two consecutive frames of a degraded sequence, the same frames of the sequence filtered by the neighborhood filter and below by the linear correction, finally the difference between the two consecutive frames of the filtered sequences. The edges of the filtered sequence by the neighborhood filter are much more irregular than the original ones. These are better regularized by the linear regression correction. The two difference images show that the oscillations from frame to frame near the edges are less noticeable in the filtered sequence by the linear regression neighborhood filter.